DOCUMENT RESUME

ED 368 077 EA 025 720

AUTHOR Bingham, C. Steven
TITLE Class Size as an Early Intervention Strategy in White-Minority Achievement Gap Reduction.
PUB DATE Feb 94
NOTE 51p.; Paper presented at the Annual Meeting of the American Association of School Administrators (San Francisco, CA, February 1994).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Academic Achievement; Achievement; *Achievement Rating; Black Students; *Class Size; Cost Effectiveness; *Early Intervention; *Minority Group Children; Primary Education; *Small Classes; *White Students

ABSTRACT

A reanalysis of norm- and criterion-referenced achievement test data from Tennessee's Student Teacher Achievement Ration Project builds on research utilizing class size and student achievement as an independent and a dependent variable, respectively. The study examined patterns of achievement among minority and white students in small, regular, and regular-with-aide class-size treatments over varying lengths of time and at varying startup years, grades K, 1, 2, and 3. The major differences between the minority and white patterns was the magnitude of the between-treatment subsample means. Differences for white subsamples tended to be compressed relative to differences between minority subsamples; that is, class size appeared to make a bigger difference for minority students than for white students. The findings of this reanalysis support the following conclusions: (1) small class size may be an effective strategy for reducing the achievement gap between whites and minorities in the primary school years; (2) benefits accrue initially in greater measure to minorities than to whites; (3) optimally beneficial treatment begins no later than grade 1; (4) optimally beneficial treatment lasts at least 2 years; (5) the major benefit of small classes for minorities appears to be in preventing rather than remediating achievement disadvantages; (6) the differential effect of small class size for minorities appears to "fade" following 2 years of treatment; and (7) the "fade" phenomenon of early treatments can be neither confirmed nor contradicted by reanalysis. (Author/KDP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 368 077

CLASS SIZE AS AN EARLY INTERVENTION STRATEGY IN WHITE-MINORITY ACHIEVEMENT GAP REDUCTION*

C. Steven Bingham
Asheboro (NC) City Schools

Paper Presented at the
Conference-within-a-Conference
American Association of School Administrators
San Francisco, California
February 12, 1994

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

C.S. Bingham

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

*For their contributions to this study, the author thanks Student Teacher Achievement Ratio (STAR) Project researchers C. M. Achilles of the University of North Carolina at Greensboro and B. Nye and D. Fulton of the Tennessee Center of Excellence for Research in Basic Skills, Tennessee State University, Nashville, TN. (Results expressed in this study do not necessarily reflect the position of the Center or its staff.) Thanks also to personnel in the Knox County, TN schools for confirming grade retention within a portion of the STAR data.

EA 025726

BEST COPY AVAILABLE

Class Size as an Early Intervention Strategy In White-Minority Achievement Gap Reduction*

C. Steven Bingham

Abstract

Building on research utilizing class size and student achievement as an independent and a dependent variable, respectively, the author examined patterns of achievement among minority and white students in small (S), regular (R), and regular-with-aide (RA) class-size treatments over varying lengths of time and at varying start-up years, grades K, 1, 2, and 3. In a re-analysis of norm- (NRT) and criterion-referenced (CRT) achievement test data from Tennessee's Student Teacher Achievement Ratio (STAR) Project, the writer reports the following results: (1) Minority and white achievement for all years and class-sizes indicated that subsamples in (S) tended to obtain higher NRT and CRT mean scaled scores in reading and mathematics than either (R) or (RA) subsamples. Analyses consistently revealed a pattern of comparatively high 1 year mean scores with diminishing 2 years and 3 years scores followed by the greatest achievement in 4 years subsamples. The major difference between the minority and white pattern was the magnitude of the between-treatment subsample means: differences for white subsamples tended to be compressed relative to differences between minority subsamples, that is, class-size appeared to make a bigger difference for minority students than for white students. Differences between minority and white achievement for all years and class-sizes tended to be smallest for the 2 years and largest for the 3 years time in treatment. Achievement differences between minority (S) classes and white (R) classes for all years revealed the smallest gaps at grade K and grade 1, 2 years. Compared to white-minority differences displayed within any one class-size, the gaps that resulted from white (R) and minority (S) were smaller on every achievement measure and at every grade level. Achievement differences between white (R) and minority (R) and (S) cohorts that matriculated synchronistically with the four years of STAR class-size treatment were smallest after one year (CRT outcomes) or two years (NRT outcomes).

The findings of this re-analysis support the following conclusions: (1) Small class-size may be an effective white-minority achievement gap reduction strategy in the primary school years; (2) Benefits accrue initially in greater measure to minorities than to whites; (3) Optimally beneficial treatment begins no later than grade 1 (preferably in K); (4) Optimally beneficial treatment lasts at least two years; (5) The major benefit of small-classes for minorities appears to be in preventing rather than remediating achievement disadvantagement; (6) The differential effect of small class-size for minorities appears to "fade" following two years treatment or in grade 3; and (7) As it applies uniquely to small class-size, the "fade" phenomenon of early treatments can be neither confirmed nor contradicted by this re-analysis.

---

*For their contributions to this study, the author thanks Student Teacher Achievement Ratio (STAR) Project researchers C. M. Achilles of the University of North Carolina at Greensboro and B. Nye and D. Fulton of the Tennessee Center of Excellence for Research in Basic Skills, Tennessee State University, Nashville, TN. (Results expressed in this study do not necessarily reflect the position of the Center or its staff.) Thanks also to personnel in the Knox County, TN schools for confirming grade retention within a portion of the STAR data.

Class Size as an Early Intervention Strategy in White-Minority Achievement Gap Reduction

The Problem

One of the greatest challenges facing U. S. education policymakers and practitioners today is reducing differences in achievement among students of different ethnic groups. Despite research suggesting that schools can educate students successfully without regard to students' categorical membership (e.g., Edmonds, 1979; Lezotte, 1989), many schools fail to do so. Specifically, students of color and poverty tend to achieve less than their white and economically advantaged peers. Forty years after Brown vs. Topeka and nearly 30 years after the Elementary and Secondary Education Act of 1965 pledged our national commitment to equity and excellence in education, the white-minority gap is still unacceptably large and troubling (Jennings, 1992; Levin, 1990; Mullis, Owens, & Phillips, 1990).

The size and nature of the achievement gap depend on a combination of the indicator, the location, and the minority group in question. Data compiled by the Sandia National Laboratory (1991) are illuminative: The Scholastic Aptitude Test (SAT) 1991 national mean score for Asian-Americans, for example, was actually higher than that of their white classmates. Similarly, retention and college attendance data indicated that there is, in fact, no achievement gap with respect to the Asian-American minority student. Black and Hispanic student data, on the other hand, demonstrated wide disparities. Using again the 1991 SAT as an indicator, the mean score for black students was approximately 200 points below the white student mean while Hispanic students experienced, on average, a 130 point disadvantage. Similarly, black and Hispanic minorities were shown to drop out of school at proportionately higher rates than their white classmates. The data indicated that nearly 80% of white students complete high

school "on time," and about 88% do so by age 25. Alternately, only 70% of black students and 50% of Hispanics graduate "on time." By age 25, the graduation rate for blacks has climbed to about 82% (still 6% lower than whites), while only 60% of Hispanics have acquired their diplomas. When location is factored in, however, the data showed that the major gaps in white-minority achievement were occuring in the large urban school districts and could be as readily explained as a function of poverty as of ethnicity.

Other databases have shown equally severe educational achievement disadvantages of minority children. The Southern Regional Education Board (SREB, 1986) reported that minority students are often "two to three or more grade levels behind on achievement measures. On National Assessment Tests, the reading achievement levels of Black eleventh grade students is basically the same as for White seventh graders (p. iv)." Results of the National Assessment of Educational Progress (NAEP) reading and mathematics tests for students grades 4, 8, and 12 (Mullis, Owens, & Phillips, 1990) demonstrated large white-minority differences in mean scores at every grade level and on both tests. A preliminary report of the 1992 NAEP assessment stated: "Although average performance for white students increased at each grade level . . . achievement for black and Hispanic students increased only at grade 12, and the gaps between whites' and blacks' performance remained substantial (Rothman, 1993a, p. 23)." Education Week later reported: "Wide Racial Gap Found on Open-Ended Math Items (Rothman, 1993b, p. 18);" and "Students' Reading Skills Fall Short, NAEP Data Find Achievement Gap for Blacks Remains Wide (Viadero, 1993, p. 1)."

Differences abound, too, in terms of placement in special educational programs (Hathaway, 1990, p. 224): black students are approximately three times as likely to be in a class for the educable mentally retarded but only half as likely to be in a class for the gifted and talented; only about one-third of the estimated 2.7 million limited English speaking proficient

(LEP) students aged 5 to 14 receive special help commensurate with their linguistic needs; black and Hispanic students are two to three times as likely to be suspended or expelled and only half as likely to be enrolled in courses that lead to college preparation. Consequently, black and Hispanic students continue to be significantly underrepresented in university degree programs (Marks, 1985; Sandia National Laboratories, 1991).

Black males appear to fare particularly poorly vis-a-vis their white classmates. Likened to an "endangered species" (Wright, 1992), black men, who make up just six percent of the U. S. population, are now three percent of college student enrollment and 47% of America's prison population (Hodgkinson, 1992, p. 3). Since 82% of America's prisoners, each of whom costs about $22,500 annually, are high school dropouts (Hodgkinson, 1992, p. 3), one need not be a rocket scientist to conclude that if schools could but improve the achievement of the black male subgroup alone, the savings to society may be profound.

In summary, whether one considers scores on norm-referenced standardized tests (NRTs) of achievement, dropout data, placement in special educational programs, or post-secondary school status, minorities lag behind their white counterparts in educational achievement. Accordingly, education policymakers must seek a reasonable option for reducing this disparity that is at once widely available, cost efficient, easily implemented and operational as a public school option. Because evidence suggests that without intervention, the achievement gap for any one minority student may increase dramatically over the course of the school career (Haycock, 1990), education researchers are challenged to investigate early intervention strategies which promise to reduce the gap. Among the fruits of such an investigation might be an enhanced understanding of all variables involved, including the comparative effects of duration and incipience of treatment (x) on student achievement. Simply stated, policy must be guided by research that answers the question: If x, how much and when?

Purpose and Significance of the Study

Results of the Student/Teacher Achievement Ratio (STAR) Project (Word, Johnson, Pate-ABain, Fulton, Zaharias, Achilles, Lintz, Folger, & Breda, 1990) and its progeny, the Lasting Benefits Study (LBS) (Nye, Achilles, Zaharias, Fulton, & Wallenhorst, 1992a) and Project CHALLENGE (Nye et al., 1992b), suggested that small class-size in the primary grades (K-3) may be effective in closing the white-minority achievement gap. Funded by the Tennessee legislature, STAR was a large-scale experiment in which kindergarten students and teachers were randomly assigned to small and large classes within each participating school. Students remained in these class-types for four years. At the end of each year, they were measured in reading and mathematics by standardized norm- and curriculum (criterion)- based tests. The results indicated that (a) a significant benefit accrued to students in reduced-size classes in both subject areas and (b) there was evidence to suggest that minority students in small classes outperformed their peers in kindergarten classes of regular size and also gained more in reading outcomes the second and subsequent years (Finn & Achilles, 1990). Summarizing the pertinent findings, Finn and Achilles asserted:

> In addition to an overall class-size effect, there is strong indication that the performance of minority students is enhanced in the small-class setting. This important outcome is statistically confirmed only in inner-city and suburban areas, but the same trend is seen in urban and rural schools as well. Also, minority students in the longitudinal subsample experienced greater relative growth than white students in the second year of small-class participation (p. 574).

Despite the minority student achievement benefits attributed to small class-size, however, the variables of (a) duration of (i.e., time spent in) "treatment" and (b) incipience of (i.e., year of first exposure to) "treatment" in a controlled experimental study remained unnaccounted for. (Neither LBS nor Project CHALLENGE are experiments.) Essentially missing from the knowledge base, then, was a quantative description of student achievement as a function

of length of time in and school year(s) of exposure to class-size treatments. Accordingly, the primary purpose of this study was to examine patterns of achievement among white and minority students in small (S), regular (R), and regular-with-aide (RA) class-size treatments over varying lengths of time and at varying start-up years, grades K, 1, 2, and 3. The researcher was guided by STAR, LBS, and CHALLENGE findings (Achilles, Nye, Zaharias, Fulton, & Bingham, 1993) which suggested that (a) earlier intervention may be more effective than later intervention, (b) the benefits of intervention may be cumulative, and (c) small class-size treatment may be more effective in preventing than in remediating the white-minority achievement gap. In addition to enhancing education researchers' understanding of the interaction of ethnicity and the duration and incipience of class-size "treatment" with student achievement, the findings suggest to education policymakers guidelines for effective implementation of small class-size as a cost efficient public school option.

Methodology

Commensurate with decision-oriented educational research and the use of existing databases (Cooley & Bickle, 1986; Stakenas, 1989), the present study was a re-analysis of data collected in the Student Teacher Achievement Ratio (STAR) Project (Word et al., 1990). Accordingly, the (1) STAR subjects, (2) STAR instruments and procedures, and (3) procedures employed in the present re-analysis will be discussed.

STAR Subjects

The Project STAR design provided for four years of randomly assigned small (S), regular (R), and regular-with-aide (RA) classroom "treatments" for a single cohort of approximately 7,000 student subjects who began Tennessee public school kindergarten in 1985 and who completed third grade in 1989. Ideally, STAR researchers intended for students assigned to one of the three class types as kindergarteners to persist in that class type all four

years; however, as new students moved into the attendance zone of an experimental school, they, too, were randomly assigned to one of the three conditions. Similarly, STAR researchers encountered subject attrition through the moving away of students from experimental schools and classrooms. Consequently, not all STAR students received four years of treatment. By the final year of the project, about one-third of the students had been in the same class type all four years; the remaining two-thirds were replacements and additions.

STAR Instruments and Procedures

Although STAR researchers collected data using instruments specifically selected for Project STAR (e.g., the Self-Concept and Motivation Inventory), the primary student achievement outcome measures were obtained from instruments mandated by the the TN Department of Education Testing Program. Both norm-referenced tests (NRT's) and criterion-referenced tests (CRT's) were included. The NRT's were nationally-normed, published instruments, while the CRT's were developed by the TN Department of Education. Commensurate with the state testing program, the instruments and testing procedures employed were a function of whether the sample was in kindergarten or in grades 1- 3.

Kindergarten. In the kindergarten sample, only a NRT, the Stanford Early School Achievement Test (SESAT) Form 2 (The Psychological Corporation, 1985), was administered. Developed as a downward extension of the Stanford Achievement Test (SAT) Series, the SESAT is a group test intended to measure school achievement in grades K.5-1.5. Scores were obtained in (1) Sounds and Letters, (2) Word Reading, (3) Total Reading, (4) Mathematics, (5) Listening to Words and Stories, (6) Total for Basic Battery, (7) Environment, and (8) Total for Complete Battery. Although validity coefficients (Person Product-Moment) between the SESAT and the Otis-Lennon School Ability Test (.81 for the complete battery) and reliablility coefficients (Kuder-Richardson 20) between the SESAT and the SAT Primary 1 instrument

(.45 to .52 on specific subscales) are modest, Ackerman (1989) suggests, "At a level of aggregation that considers intact classes as a whole in comparison to state and national norms, the test may provide useful information (p. 866)." For the purposes of this study, only the Total Reading and Mathematics scores were analyzed. According to the Technical Data Report (The Psychological Corporation, 1985) compiled by the test publisher, KR-20 r's were .93 and .81 for the total reading and total math respectively (p. 32). The SESAT was administered during the spring of the kindergarten year under controlled testing conditions including the use of trained substitute teachers as proctors.

Grades 1-3. Beginning in grade1, the Stanford Achievement Test (SAT) Primary 1 was employed as the NRT instrument. The SAT is a nationally normed group test intended to measure "the important learning outcomes of the school curriculum (The Psychological Corporation, 1985)." Scores for the Primary 1 form (administered in grades 1 and 2) were obtained in (1) Word Study Skills, (2) Word Reading, (3) Reading Comprehension, (4) Total Reading, (5) Concepts of Number, (6) Mathematics Computation, (7) Mathematics Applications, (8) Total Mathematics, (9) Language, (10) Spelling, (11) Environment, (12) Listening, (13) Basic Battery, and (14) Complete Battery. Primary 2 form, administered in grade 3, is identical to Primary 1 except for substituting "Reading Vocabulary" for "Word Reading." Kuder-Richardson (KR) 20 reliability coefficients have been obtained for each test and subtest for each form and level. These range from .85 to over .90. Alternative form reliability coefficients tend to be slightly lower than the corresponding KR-20 coefficients, but almost all are .80 or higher (Carpenter, 1989). The present study utilized only the Total Reading and Total Mathematics subscores as student achievement NRT outcome measures.

Beginning in grade 1, the TN Department of Education Testing Program provided for the use of state-developed curriculum-referenced (criterion-referenced) tests locally known as Basic

Skills First (BSF) tests (TN Department of Education, 1987). These tests were created from well-specified lists of objectives in reading and mathematics at each grade level, and can be scored either as the total number of items answered correctly, or as pass-fail. A student passes if he/she masters 80% of the objectives covered by the test items. The present study used the pass-fail method in the reported analyses. No reliablity data are available for the BSF tests. The SAT and BSF tests were administered during the spring of the grades 1, 2, and 3 years under controlled conditions.

Procedures Employed in the Present Re-Analysis

To examine the absolute and comparative achievement patterns of minority versus white students in small (S), regular (R), and regular-with-aide (RA) class-sizes as a function of time in "treatment" and beginning schoolyear of treatment required: (1) identifying the appropriate subsample of students in each grade and class-size and for each achievement measure; (2) calculating the minority and white subsamples' mean scaled scores and standard deviations for each achievement measure by length of time in each class-size treatment at each grade; (3) tabulating the minority and white subsamples' mean scaled scores for each achievement measure by length of time in each class-size treatment at each grade; (4) tabulating a white-minority achievement gap as shown by differences in mean scaled scores of each achievement measure by grade by length of time in each class-size treatment; (5) tabulating achievement differences between minority small-class and white regular-class subsamples by grade by length of time in class; (6) calculating "effect sizes" to measure the "educational importance" of selected mean score achievement differences between subsamples; and (7) charting histograms to compare the achievement of white (R) and minority (R) and (S) class-sizes. Regarding procedure (6), Glass and Hopkins (1984, p. 54) have explained that "educational importance" or "effect size" (ES) can be mathematically represented as the

quotient of the difference (gain or loss) between any two means and the standard deviation of their combined distributions. It is expressed as a proportion of 1.00 sigma.

Results

Minority Student Achievement

Table 1 shows the results for (S) exposure. Except for the 4 years subsample, both SAT (the NRT component) mean scaled scores of minority subsamples which spent 1 year in (S) treatment exceed that of subsamples exposed 2 years or 3 years. Similarly, the BSF (the CRT component) measures of (S) minority students are, except for the subsample that began in K and continued through grade 3, greater for each 1 year subsample than for corresponding multiple year subsamples. In fact, except for the subsample that spent 4 years in (S), increased time in treatment appears to be associated with lower achievement. On average, both SAT and BSF mean differences between 1 year, 2 years, and 3 years exposed subsamples were moderately small with effect sizes (ES) < .20 sigma. At 18.73 points (.49 sigma), the point of maximal difference was the SAT Math mean scores between the 3 years and 4 years subsamples. (Subsample n's and SD's are reported in Appendix A and Appendix B, respectively.)

Table 2 shows the results for minorities exposed to (R). With one exception (BSF Math, 2 years), the same pattern of achievement attained by (S) subsamples was shown by (R) subsamples. Although displaying a wide range of differences across any one grade, the achievement for 1 year subsamples tended to be greater than that of multiple year students. At 16.88 points (.43 sigma), the point of maximal difference was again the SAT Math mean scores between the 3 years and 4 years subsamples.

Table 3 shows the results for minorities exposed to (RA). As in the (S) and (R) subsamples, 1 year achievement measures uniformly exceeded 2 years measures; however, 3 years subsample mean scores were greater than 1 year or 2 years measures while 4 years

Table 1

Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Small Class-Size

Achievement Measures	Grade	Length of Time in Small Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	439.12	-----	-----	-----
	1	510.18	495.35	-----	-----
	2	571.51	569.18	561.37	-----
	3	607.79	606.19	594.55	612.78
BSF Reading	K	-----	-----	-----	-----
	1	27.22	24.83	-----	-----
	2	38.64	37.16	35.17	-----
	3	31.17	30.71	28.03	32.11
SAT Math	K	479.56	-----	-----	-----
	1	521.49	503.40	-----	-----
	2	568.91	566.93	556.46	-----
	3	608.79	604.39	594.90	613.63
BSF Math	K	-----	-----	-----	-----
	1	38.96	35.86	-----	-----
	2	51.90	51.90	48.49	-----
	3	48.00	47.62	44.84	48.85

Table 2

Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size

Achievement Measures	Grade	Length of Time in Regular Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	425.60	_____	_____	_____
	1	491.36	468.23	_____	_____
	2	558.82	546.73	555.95	_____
	3	595.81	590.18	592.23	607.89
BSF Reading	K	_____	_____	_____	_____
	1	24.35	20.41	_____	_____
	2	35.84	31.67	35.84	_____
	3	30.01	29.24	29.04	31.87
SAT Math	K	472.53	_____	_____	_____
	1	508.39	484.98	_____	_____
	2	559.20	542.68	558.21	_____
	3	600.62	599.73	593.00	609.88
BSF Math	K	_____	_____	_____	_____
	1	36.45	31.92	_____	_____
	2	49.67	46.71	48.56	_____
	3	46.08	45.07	44.78	48.24

Table 3

Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular/Aide Class-Size

Achievement Measures	Grade	Length of Time in Regular/Aide Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	428.61	_____	_____	_____
	1	493.18	475.33	_____	_____
	2	559.72	542.57	556.80	_____
	3	598.26	594.57	600.46	605.18
BSF Reading	K	_____	_____	_____	_____
	1	24.83	21.93	_____	_____
	2	35.81	30.74	36.28	_____
	3	30.00	29.84	29.97	31.16
SAT Math	K	469.64	_____	_____	_____
	1	509.61	494.58	_____	_____
	2	557.51	531.85	557.15	_____
	3	598.98	595.41	604.80	605.51
BSF Math	K	_____	_____	_____	_____
	1	36.81	33.13	_____	_____
	2	49.91	44.42	49.50	_____
	3	46.21	45.75	47.28	47.46

means exceeded 3 years means. After 1 year, then, there appears to be a positive linear relationship between length of time in (RA) and achievement. The average ES between the lowest SAT scores in the 2 years subsamples and the highest scores in the 4 years subsamples was about .32 sigma. Both BSF outcomes in the (RA) 3 years exposure at grade 2 subsample exceeded that of students who had been similarly exposed at grade 3 for all 4 years by about 5.00 percentage points (.67 sigma).

White Student Achievement

Table 4 shows the results of white student subsamples exposed to (S) treatment. For both SAT and BSF measures, 1 year subsamples uniformly achieved greater than 2 years and 3 years subsamples. White students who began (S) treatment in K and who continued all 4 years, however, showed the highest levels of achievement. The ES differences between the 4 years subsamples and the 1 year subsamples (about .16 sigma on the SAT measures and .08 sigma on the BSF measures) were small. The differences between the 3 years and 4 years subsamples, however, increase to .38 sigma and .30 sigma for the SAT and BSF respectively. Obvious in its singularity, the grade 2, 3 years subsample outscored the grade 3, 4 years subsample by 3.67 percentage points (.63 sigma) on the BSF reading measure.

Table 5 shows the results of white student subsamples exposed to (R). The pattern of relatively high 1 year mean scores with diminishing 2 years and 3 years scores, followed by the greatest achievement from 4 years subsamples, recurs in these data. Achievement ES differences between highest and lowest mean scores average about .43 on the SAT. Student subsamples in grade 2, regardless of length of time spent in (R), outscore student subsamples exposed 4 years by an average of more than 4.00 percentage points (.67 sigma) on BSF measures. The one exception is at grade 2, 1 year.

Table 4

White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Small Class-Size

Achievement Measures	Grade	Length of Time in Small Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	443.70	-----	-----	-----
	1	539.49	513.31	-----	-----
	2	599.95	570.18	578.14	-----
	3	627.15	619.10	619.07	633.43
BSF Reading	K	-----	-----	-----	-----
	1	27.81	24.55	-----	-----
	2	41.49	35.42	38.30	-----
	3	34.04	33.58	33.33	34.63
SAT Math	K	496.20	-----	-----	-----
	1	546.67	523.72	-----	-----
	2	595.46	576.87	575.46	-----
	3	629.25	621.32	620.35	635.12
BSF Math	K	-----	-----	-----	-----
	1	40.45	36.34	-----	-----
	2	54.15	51.45	51.15	-----
	3	53.18	52.55	51.86	54.01

Table 5

White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size

Achievement Measures	Grade	Length of Time in Regular Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	439.07	_____	_____	_____
	1	527.31	501.43	_____	_____
	2	591.97	579.68	578.28	_____
	3	621.59	618.98	610.61	627.51
BSF Reading	K	_____	_____	_____	_____
	1	26.78	24.16	_____	_____
	2	40.64	38.10	38.66	_____
	3	33.35	33.29	32.09	34.06
SAT Math	K	488.34	_____	_____	_____
	1	535.53	514.83	_____	_____
	2	588.63	578.50	572.95	_____
	3	624.26	623.73	615.92	630.43
BSF Math	K	_____	_____	_____	_____
	1	39.78	36.15	_____	_____
	2	53.23	50.57	50.75	_____
	3	52.16	52.26	50.17	53.29

Table 6 shows the results of white student subsamples treated in (RA). The 1 year subsamples again demonstrate consistently higher levels of achievement than either 2 years or 3 years subsamples. Achievement measures for 4 years subsamples, however, remain uniformly highest. The average ES difference between the lowest and highest SAT mean scaled scores is about .44 sigma. BSF Reading measures at grade 2, regardless of time spent in regular-with-aide class-size, are consistently higher than BSF measures on student subsamples exposed 4 years.

Minority Compared with White Student Achievement

Having examined the patterns of achievement for minority students and white students separately , the comparative achievement, or the achievement gap, of white and minority students as a function of class-size type, duration, and incipience of treatment. As in the previous analyses, tabulations are made of each class-size for white and minority subsamples separately (Tables 7-9). However, to determine the possible effect of (S) in preventing the achievement gap, the researcher also cross-tabulated minority (S) and white (R) achievement differences (Table 10). To compare graphically the K, K+1, K+1+2, and K+1+2+3 achievement differences for white (R) and minority (R) and minority (S) subsamples, the researcher charted histograms, one for each achievement measure (Figures 1-4).

Table 7 shows the white-minority achievement gap by grade and length of time in (S). With one exception (BSF Reading, grade 3, 4 years), the 2 years gaps are smallest and the 3 years gaps largest. In grade 3, the second smallest gaps vascillate between 1 year and 4 years length of time spent in (S). The sizes of the gaps are large, ranging from 27.44 points between the grade 1,1 year and 2 years subsamples on the SAT Reading measure to 6.57 percentage points between the grade 2, 2 years and grade 3, 3 years subsamples on the BSF Math measure.

Table 6

White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular/Aide Class-Size

Achievement Measures	Grade	Length of Time in Regular/Aide Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	438.88	_____	_____	_____
	1	533.38	504.20	_____	_____
	2	595.38	580.77	582.56	_____
	3	621.21	615.50	612.94	628.09
BSF Reading	K	_____	_____	_____	_____
	1	27.03	23.68	_____	_____
	2	40.86	37.69	38.81	_____
	3	33.17	32.28	32.00	34.14
SAT Math	K	489.53	_____	_____	_____
	1	537.95	517.48	_____	_____
	2	590.32	575.62	584.91	_____
	3	624.10	618.22	614.63	631.91
BSF Math	K	_____	_____	_____	_____
	1	39.91	36.95	_____	_____
	2	53.71	50.57	51.80	_____
	3	51.96	51.45	49.84	53.31

Table 7

The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Small Class-Size

Achievement Measures	Grade	Length of Time in Small Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	9.91	-----	-----	-----
	1	29.31	17.96	-----	-----
	2	28.44	1.00	16.77	-----
	3	19.36	12.91	24.52	20.65
BSF Reading	K	-----	-----	-----	-----
	1	0.59	0.28	-----	-----
	2	2.85	1.74	3.13	-----
	3	2.87	2.87	5.30	2.52
SAT Math	K	16.64	-----	-----	-----
	1	25.18	20.32	-----	-----
	2	26.55	9.94	19.00	-----
	3	20.46	16.93	25.45	21.49
BSF Math	K	-----	-----	-----	-----
	1	1.49	0.48	-----	-----
	2	2.25	0.45	2.66	-----
	3	5.18	4.93	7.02	5.16

Table 8 shows the white-minority achievement gap by grade and length of time in (R). Except for the apparent moderate negative linear relationship between the achievement gap and length of time spent in (R) suggested by the grade 2 and grade 3 outcomes, no clear patterns seem to emerge. Even in the case of the latter, three out of the four achievement measures register a difference score gain after 2 years treatment. In short, the gap between both SAT and BSF measures seems to vary almost randomly.

Table 9 shows the white-minority achievement gap obtained from (RA) subsamples. Clearly, the grade 3, 3 years length of time in (RA) subsamples mean differences were smallest. The average difference between the highest and lowest gap score for the SAT measures was more than 22.00 points, while the average percentage passing difference on the BSF was more than 2.00 points. For each measure, the grade 3, 4 years subsamples gap was either largest or next-to-largest.

Table 10 shows minority (S) and white (R) achievement differences. On every achievement measure, the gaps are smallest at grade K (SAT Reading and Math) or grade 1, 2 years (BSF Reading and Math). Except for SAT Reading, grade 2, 2 years and SAT Math, grade 2, 1 year, the grade 3, 3 years gaps are largest. Uniformly, the grade 3, 4 years achievement gaps are smaller than the grade 3, 3 years gaps and, except for SAT Reading, smaller than the 2 years gaps. The sizes of these gaps are consistent with the other analyses, ranging up to .16 standard deviation and over 2.00 percentage points. Additional results may be obtained by comparing Table 10 with the other gap tables (Tables 7-9) showing the white-minority differences for any one class-size treatment. On every achievement measure and at every grade level, the obtained differences are smaller, sometimes by as much as 23.00 points.

Table 8

The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size

Achievement Measures	Grade	Length of Time in Regular Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	11.10	_____	_____	_____
	1	35.95	33.20	_____	_____
	2	33.15	32.95	22.33	_____
	3	25.78	28.80	18.38	19.62
BSF Reading	K	_____	_____	_____	_____
	1	2.43	3.75	_____	_____
	2	4.80	6.43	2.82	_____
	3	3.34	4.05	3.05	2.19
SAT Math	K	10.57	_____	_____	_____
	1	27.14	29.85	_____	_____
	2	29.43	35.82	14.74	_____
	3	23.64	24.00	22.92	20.55
BSF Math	K	_____	_____	_____	_____
	1	3.33	4.23	_____	_____
	2	3.56	3.86	2.19	_____
	3	6.08	7.19	5.39	5.05

Table 9

The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Reg/Aide Class-Size

Achievement Measures		Length of Time in Regular/Aide Class-size			
	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	10.27	-----	-----	-----
	1	40.20	28.87	-----	-----
	2	35.16	38.20	25.76	-----
	3	22.95	20.93	12.48	22.91
BSF Reading	K	-----	-----	-----	-----
	1	2.20	1.75	-----	-----
	2	5.05	6.95	2.53	-----
	3	3.17	2.44	2.03	2.98
SAT Math	K	19.89	-----	-----	-----
	1	28.34	22.90	-----	-----
	2	32.81	43.77	27.76	-----
	3	25.12	22.81	9.83	26.40
BSF Math	K	-----	-----	-----	-----
	1	3.10	3.82	-----	-----
	2	3.80	6.15	2.30	-----
	3	5.75	5.70	2.56	5.85

Table 10

Minority Small-Class and White Regular-Class Subsample Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Class

Achievement Measures	Grade	Length of Time in Class: 1 year	2 years	3 years	4 years
SAT Reading	K	5.35	-----	-----	-----
	1	17.13	6.08	-----	-----
	2	20.46	10.50	16.91	-----
	3	13.80	12.79	16.06	14.73
BSF Reading	K	-----	-----	-----	-----
	1	0.44	0.67	-----	-----
	2	2.00	0.94	3.49	-----
	3	2.18	2.58	4.06	1.95
SAT Math	K	8.46	-----	-----	-----
	1	14.04	11.43	-----	-----
	2	19.72	11.57	16.49	-----
	3	15.47	19.34	20.83	16.80
BSF Math	K	-----	-----	-----	-----
	1	0.82	0.29	-----	-----
	2	1.33	1.33	2.26	-----
	3	4.16	4.64	5.33	4.44

Figures 1-4 graphically illustrate the comparative achievement of white (R) and minority (R) and minority (S) for each achievement measure across one year (K), two years (K+1), three years (K+1+2), and four years (K+1+2+3). Figure 1 shows the results for SAT Reading. At each grade level combination, the gap between white (R) and minority (S) is narrower than that created by the two (Rs). The gap between white (R) and minority (S) achievement is narrowest at grades K and K+1 and grows increasingly larger in K+1+2 and K+1+2+3. In the last two periods, minority (S) and minority (R) achievement is nearly identical. Overall year-to-year achievement gains, however, are substantially higher for all subsamples.

Figure 2 shows the results for BSF Reading. (No BSF tests were administered in grade K.) In grade 1, the minority (S) subsample outscored both white (R) and minority (R) subsamples. In subsequent grades, minority (S) achievement was virtually the same as minority (R) achievement, while white (R) achievement was comparatively and uniformly higher. For all subsamples, initial surges in achievement from 1 to 1+2 were followed by declines in 1+2+3.

Figure 3 shows the results for SAT Math. This histogram is almost identical to that for SAT Reading (Figure 1). The greatest comparative differences occur in K+1. The difference in minority (R) and (S) is neglible for both K+1+2 and K+1+2+3.

Figure 4 shows the results for BSF Math. As in the case of the histograms for the SAT measures, the BSF histograms are remarkably similar. Initial differences (grade 1) are maximal. In subsequent periods, the white (R) subsamples achieve passing rates several percentage points higher than minority (R) or (S) subsamples. Differences between both minority subsamples are but a few hundreths of a percentage point.

Figure 1

Comparison of SAT Reading Mean Scaled Scores for White (R) and Minority (R) and (S) Subsamples That Began in K and Matriculated All Four Years of STAR

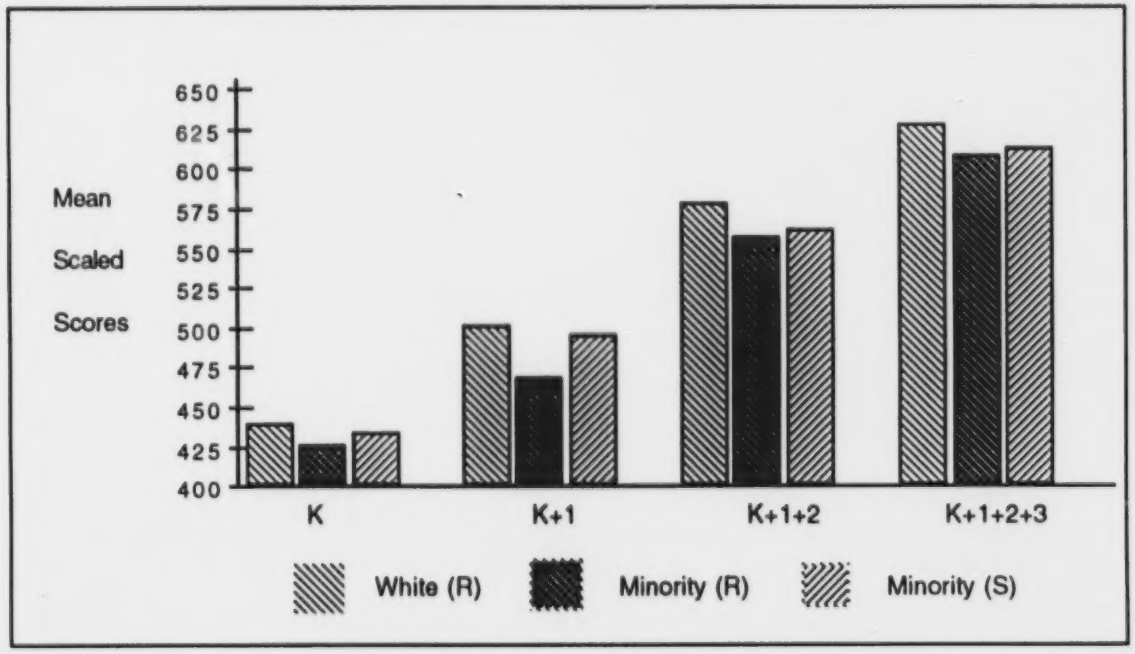


Figure 2

Comparison of BSF Reading Mean Percentage Passing for White (R) and Minority (R) and (S) Subsamples That Began in K and Matriculated All Four Years of STAR

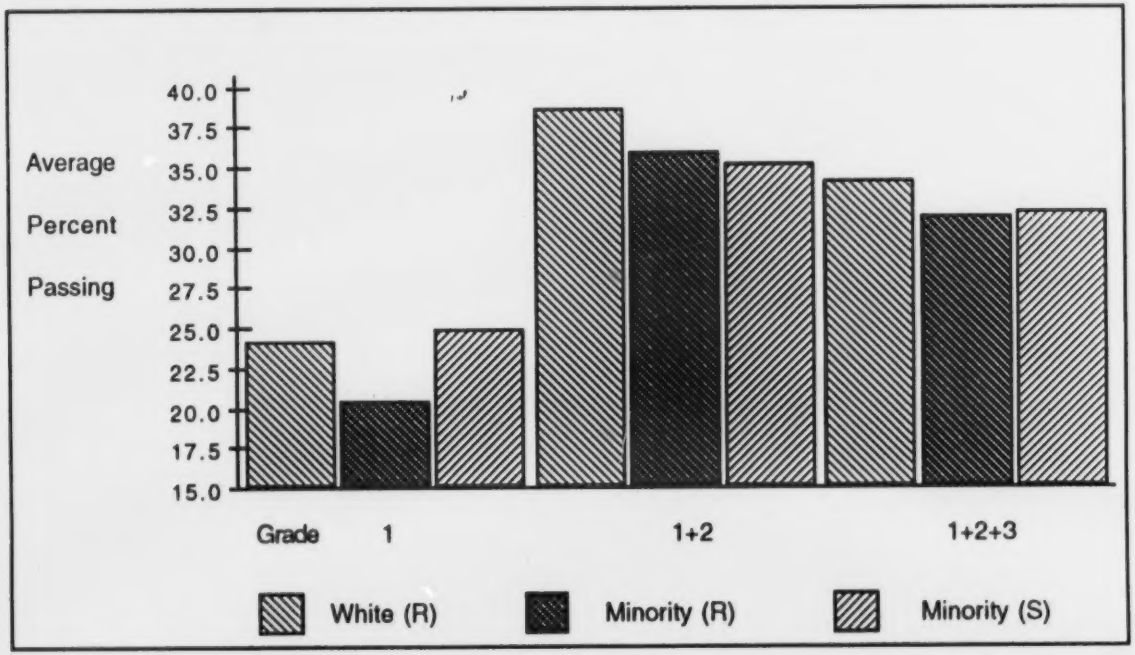


Figure 3

Comparison of SAT Math Mean Scaled Scores for White (R) and Minority (R) and (S) Subsamples That Began in K and Matriculated All Four Years of STAR

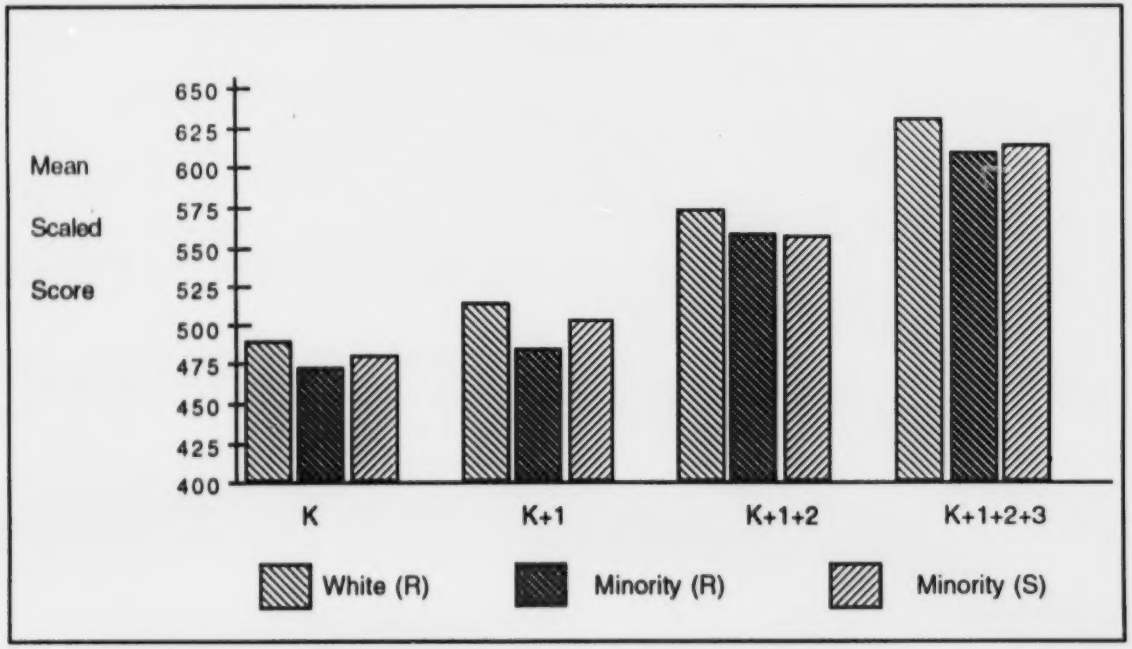
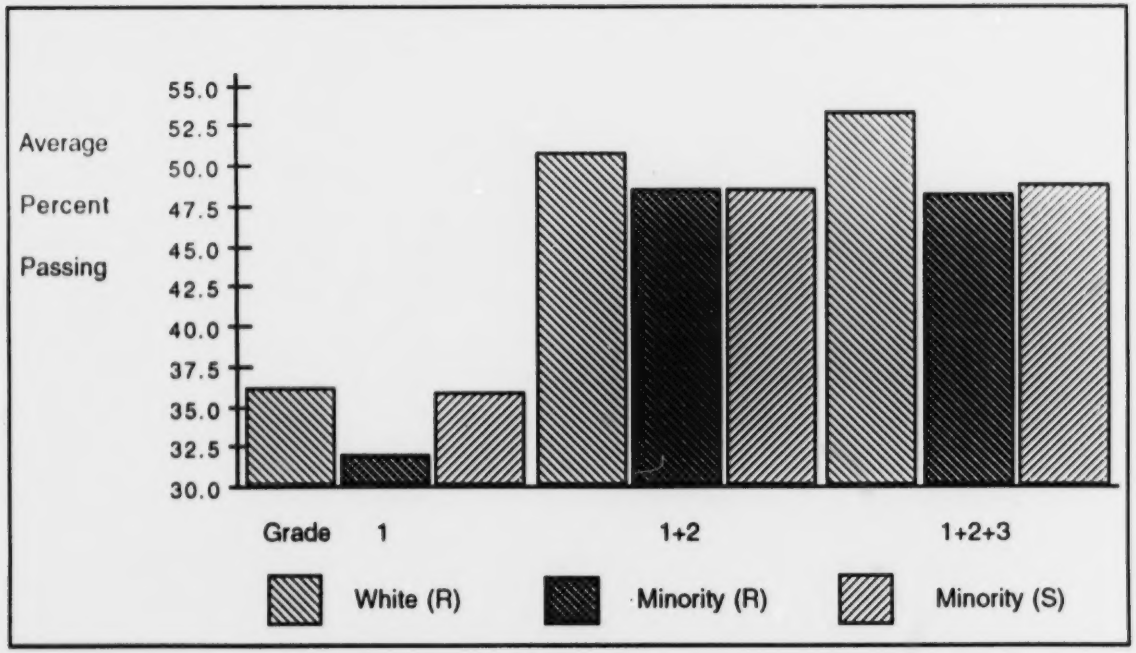


Figure 4

Comparison of BSF Math Mean Percentage Passing for White (R) and Minority (R) and (S) That Began in K and Matriculated All Four Years of STAR



BEST COPY AVAILABLE

Conclusions

The purpose of this investigation was to examine patterns of achievement among white and minority students in small (S), regular (R), and regular-with-aide (RA) class-size treatments over varying lengths of time and at varying start-up years, grades K, 1, 2, and 3. The investigator reasoned that "best" achievement gap reduction intervention would involve delivering treatment in the optimally effective quantity (length of time) and at the optimally effective time (start-up year). The investigator reasoned further that, to be a politically and educationally acceptable goal, small class-size (or any class-size) as a gap reduction intervention must be shown not only to benefit minorities, but not to harm whites. Consequently, the investigator examined absolute (within a group) and comparative (between the two groups) achievement for minorities and whites. To illustrate the effect of no intervention (considered in this context as R) vs. the effect of minority (S) as a gap reduction intervention strategy, the investigator extended the comparative examination to include different class-sizes within and between groups.

The Length of Time Issue

Relative to the length of time issue, the results showed that absolute achievement on NRT and CRT measures of both minorities and whites was optimal (highest) for students who had experienced all four years of small-class treatment. Moreover, minority between class-size differences were relatively larger than white between class-size differences, a fact which suggests that the effect of small classes on achievement was of greater benefit to minorities than to whites. This finding supports earlier STAR, LBS, and CHALLENGE research (Finn & Achilles, 1990; Achilles et al., 1992).

The comparative achievement analyses indicated, however, that the differential benefit for minorities did not seem to last. The white-minority achievement gap, smallest after two

years time in small-class treatment, appeared to "fade" in subsequent years. (The "fade" issue will be addressed in the next section of this paper.) The same finding was obtained in the minority small-class vs. white regular-class contrast; however, the resulting gaps were uniformly smaller on every measure and at every grade. The diminishing differential effect of small class-size for minorities was graphically apparent in the contrast of that group with white and minority regular-class students that participated in all four years of STAR.

The Start-Up Year Issue

When one considers absolute achievement, the question of optimizing the benefit of small class-size by starting treatment in one or another school year highlights an intriguing finding of the present study: Except for the cohort that started in grade K and spent four years time in treatment, one year exposed students score higher than multiple year exposed students. The answer to this anomoly may, in part, be a function of the retention-promotion practices of STAR teachers and the scores of promoted vs. retained students. The percent of grade retentions in STAR was smallest in the (S) condition (7.8 vs. 12.6 in the R and 10.8 in the RA, grade 1) Moreover, students who were promoted tended to score lower (e.g., 422 in S vs. 427 in R on SAT Reading, K). (Percentages and scores are from Word et al., p. 171.) The combination of fewer retainees and lower scores resulted in more "marginal" students matriculating on time with the cohort. Thus when tests were administered at any given grade, more academically capable and older students entering Project STAR for the first time outscored the less capable nonretained students in the original cohort. Over time, this phenomenon results in an apparent "fade" of early intervention benefits (Barnett, 1992). One year treatment means were higher than two and three year treatment means, however, not just for small-classes but for all class-sizes. A possible explanation is that increased attention was afforded STAR students (and teachers) regardless of class-size treatment. Accompanied by the possibility of higher

expectations of and for STAR participants, a Hawthorne effect could have been generated. (Additional analyses tend not to support this supposition. See Appendix C.)

The comparative achievement pattern analyses examining white vs. minority differences for each class-size found that gap scores tended to be least for the two years in small-class group. Interestingly, this finding held true whether the start-up year was K, 1, or even 2; however, the smallest gaps were registered by the groups that began treatment in grade 1. When minority small-class achievement was compared to white regular-class achievement, the gaps were shown to be least at grade K and grade 1, two years treatment.

What is the Optimal Time in Small Class-Size and When Should It Begin?

Consistent with prior STAR, LBS, and CHALLENGE research, the findings of this re-analysis of the STAR database support the following conclusions:

(1) Small class-size may be an effective white-minority achievement gap reduction strategy in the primary school years.

(2) Benefits accrue initially in greater measure to minorities than to whites.

(3) Optimally beneficial (defined as resulting in least white-minority achievement differences) treatment begins no later than grade 1 (preferably in K).

(4) Optimally beneficial treatment lasts at least two years, grade 1 and K or 1 and 2.

(5) The differential effect of small class-size for minorities appears to "fade" following two years treatment or in grade 3.

(6) As it applies uniquely to small class-size, the "fade" phenomenon in early treatments can be neither confirmed nor contradicted by this re-analysis. To the extent that STAR small-class students were less likely to be retained and that retention in grade is an indicator that a student will drop out of school (Hahn, 1987), the nonretention of minority and white students in Project STAR broadens to life-long the optimization small-class treatment.

Implications for Policy and Practice

First, although a cost-benefit analysis exceeds the scope of this study, the expense of implementing small classes in the primary years of school must be weighed against the cost of remediation (e.g., Chapter 1 reading programs) in later years, the effectiveness of which is questionable. As a report of the Perry School Project recently disclosed (Lewis, 1993), the benefits of early childhood intervention programs tend to yield lifelong dividends (e.g., higher levels of education, higher economic status, and greater social responsibility) for self and society. For preschool programs, the latest research suggests that $7.16 is the eventual savings to the public for every dollar invested. Although exact figures are impossible to estimate at this time, small class-size, too, may be a cost-efficient public school option. Absent contradictory evidence, small classes should be implemented in grades K and 1 whenever possible.

Second, education policymakers and practitioners must act in ways that reflect sensitivity to the concerns of ethnic and racial minorities. One explanation for why small class-size may be of greater initial benefit to minorities than to whites is grounded in Cultural Deprivation Theory. According to Lansa and Potter (1984), underlying Cultural Deprivation is Piaget's theory that a child who, at an early age, is deprived of appropriate environmental stimuli lacks the experiences needed for the development of intellectual skills and abilities (p. 4). To the extent that espousing Cultural Deprivation Theory may be viewed as the oppression of a minority culture by a majority one, it is ethically wrong. Gardner (1991, p. 53), for example, reports that the Kaluli people of New Guinea see babies as helpless creatures who neither understand nor are capable of speech. Rather than speaking to the children, Kaluli mothers speak "for them." Yet the family unit is strong with the generations cohabiting until death parts them. Despite the viability and grace of Kaluli practices, the tendency to label such

a culture "deprived" vis-a-vis our predominantly white, Anglo-Saxon, protestant, middle-class American culture is great. However, educators and education policymakers must not permit application of the model in everyday life to minimize or obliterate the good and moral qualities inherent in all cultures and peoples.

Within the next 20 years, Americans will recognize that as the number of children declines as a percent of the U. S. population and people of color (non-Anglos) become the majority (see Hodgkinson, 1992), we literally cannot afford to allow the achievement of the "new majority" youngsters to lag behind their white counterparts. We cannot wait 20 years to do something about it. Preliminary research suggests that small class-size in the primary years of school may be an effective strategy for not only reducing the the white-minority achievement gap, but for enhancing the achievement of all students.

References

Achilles, C. M., Nye, B. A., Zaharias, J. B., Fulton, B. D., & Wallenhorst, M. P. (1992, August). Class size is a reasonable policy alternative for educational excellence--or is it? Paper presented at the National Conference of Professors of Educational Administration, Indiana State University, Terre Haute, IN.

Achilles, C. M., Nye, B. A., Zaharias, J. B., Fulton, D., & Bingham, S. (1993, August). Prevention or remediation: Is small class a reasonable treatment for either? Paper presented at the National Conference of Professors of Educational Administration, Palm Springs, CA.

Ackerman, P. L. (1989). Review of the Stanford Early School Achievement Test, Third Edition. In J. C. Conoley and J. J. Kramer (Eds.), The tenth mental measurements yearbook (pp. 866-867), Lincoln, NE: The Buros Institute of Mental Measurements.

Barnett, S. W. (1992, Spring). Benefits of compensatory preschool education. Journal of Human Resources, 279-312.

Carpenter, D. (1989). Review of the Stanford Achievment Test, Eighth Edition. In J. C. Conoley and J. J. Kramer (Eds.), The tenth mental measurements yearbook (pp. 861-863), Lincoln, NE: The Buros Institute of Mental Measurements.

Cooley, W., & Bickel, W. (1986). Decision-oriented educational research. Boston, MS: Kluwer-Nijhoff Publishing.

Edmonds, R. (1979, September). Effective Schools for the urban poor. Educational Leadership, 37(1), 15-24.

Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. American Educational Research Journal, 27(3), 557-577.

Gardner, H. (1991). The Unschooled mind: How children think & how schools should teach. New York: Basic Books.

Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hahn, A. (1987, December). Reaching out to America's dropouts: What to do? Phi Delta Kappan, 69(4), 256-263.

Hathaway, W. (1990, October). Monitoring and improving school learning. In J. G. Bain & J. L. Herman(Eds.), Making schools work for underachieving minority students: Next steps for research,policy, and practice. New York: Greenwood Press.

Haycock, K. (1990). Equity, relevance, and will. In J. G. Bain & J. L. Herman (Eds.), Making schools work for underachieving minority students: Next steps for research, policy, and practice. New York: Greenwood Press.

Hodgkinson, H. L. (1992, October). A demographic look at tomorrow. Paper presented at the annual fall conference of Personnel Administrators of North Carolina, Raleigh, NC.

Jennings, J. M. (1992, November). Closing the achievement gap: A model for success. Paper presented at the Southern Regional Council of Educational Administrators Annual Conference, Georgia State University, Atlanta, GA.

Lansa, P. J., & Potter, J. H. (1984, January). Building a bridge to span the minority-majority achievement gap. Paper presented at the National Conference on Desegregation in Postsecondary Education, Durham, NC.

Lewis, A. C. (1993, June). The payoff from a quality preschool. Phi Delta Kappan, 74(10), 748-749.

Levin, H. (1990). The educationally disadvantaged are still among us. In J. G. Bain & J. L. Herman (Eds.), Making schools work for underachieving minority students: Next steps for research, policy, and practice. New York: Greenwood Press.

Lezotte, L. W. (1989, August). Base school improvement on what we know about effective schools. American School Board Journal, 18-20.

Marks, J. L. (1985). The enrollment of black students in higher education: Can declines be prevented? (ERIC Document Reproduction Service No. ED 264 817).

Mullis, I., Owen, E. H., & Phillips, G. W. (1990). Accelerating academic achievement--America's challenge: A summary of findings from 20 years of NAEP. (ERIC Document Reproduction Service No. ED 325 500).

Nye, B. A., Achilles, C. M., Zaharias, J. B., Fulton, B. D., & Wallenhorst, M. P. (1992a, April). Five years of small-class research: Student benefits derived from reduced student/teacher ratios. Paper presented at the American Educational Research Association, San Francisco, CA.

Nye, B. A., Achilles, C. M., Zaharias, J. B., Fulton, B. D., & Wallenhorst, M. P. (1992b). Small is far better. Paper submitted for MidSouth Educational Research Association Oustanding Paper Competition, Knoxville, TN.

Rothman, R. (1993a, January 20). Achievement in math improves, NAEP data find. Education Week, pp. 1, 23.

Rothman, R. (1993b, June 23). Wide racial gap found on open-ended math items. Education Week, p. 18.

Sandia National Laboratories. (1991). Perspectives on education in America: Annotated briefing-third draft. Systems Analysis Department, Albuquerque, NM.

Southern Regional Education Board (SREB). (1986). Measuring student achievement. Atlanta, GA: Author.

Stakenas, R. G. (1989). Thoughts on prospectuses and dissertations for the Ed.D.: Is decision-oriented research a viable option? Unpublished manuscript, Florida State University, Department of Educational Leadership, Tallahassee, FL.

The Psychological Corporation (1985). Stanford Achievement Test series. Technical data report. Orlando, FL: Harcourt, Brace, Jovanovich.

TN Department of Education (1987). STAR criterion referenced test. Manual for test administrator. Grade 1. Nashville, TN: State Education Department.

Viadero, D. (1993, September 22). Students' reading skills fall short, NAEP data find: Achievement gap for blacks remains wide. Education Week, pp. 1, 16.

Word, E., Johnston, J., Bain, H., Fulton, B., Zaharias, J., Lintz, N., Achilles, C. M., Folger, J., & Breda, C. (1990). Student/teacher achievement ratio (STAR): Tennessee's K-3 class size study. Final report and final report summary. Nashville, TN: Tennessee State Department of Education.

Wright, W. J. (1992). The endangered black male child. Educational Leadership, 49(4), 14-16.

APPENDIX A: Achievement Measures Total n of Cases by Time in Treatment by Class-Size by Race

			K	1	2	3	K+1	1+2	2+3	K+1+2	1+2+3	K+1+2+3	
	S	W	1187	1231	1193	1311	175	16	130	83	154	831	
		M	552	592	598	602	84	29	95	62	91	309	
SAT	R	W	1359	1500	1302	1176	182	40	195	79	206	553	
Read		M	646	934	746	639	145	55	142	60	116	197	
	RA	W	1357	1492	1427	1452	169	60	182	57	278	762	
		M	687	644	790	808	91	51	213	59	147	245	
	S	W		1257	1240	1269	173	19	125	84	150	804	
		M		595	636	586	83	31	95	65	89	299	
BSF	R	W		1551	1345	1173	187	41	197	74	196	547	
Read		M		932	787	645	145	58	143	64	115	199	
	RA	W		1550	1467	1408	174	62	181	59	267	739	
		M		650	816	815	90	53	214	60	146	248	
	S	W	1204	1271	1189	1338	181	40	131	83	158	847	
		M	558	596	597	599	84	28	97	61	91	306	
SAT	R	W	1370	1550	1467	1408	187	40	194	78	211	559	
Math		M	661	946	748	637	149	56	142	61	116	196	
	RA	W	1374	1570	1424	1493	180	60	185	56	289	779	
		M	703	654	788	803	91	52	213	59	147	242	
	S	W		1255	1266	1296	172	20	126	87	153	819	
		M		586	635	587	80	30	95	65	89	300	
BSF	R	W		1543	1368	1202	183	44	201	77	202	558	
Math		M		932	787	643	144	58	143	64	114	199	
	RA	W		1552	1512	1449	175	63	186	61	279	756	
		M		644	817	814	90	53	213	60	146	248	

APPENDIX B: Achievement Measures Standard Deviations by Time in Treatment by Class-Size by Race

			K	1	2	3	K+1	1+2	2+3	K+1+2	1+2+3	K+1+2+3	
	S	W	33.52	57.92	46.74	38.84	65.71	57.28	38.60	50.34	38.89	37.36	
		M	29.07	48.06	38.34	36.97	49.89	50.96	35.45	41.66	34.62	36.23	
SAT	R	W	31.60	54.13	45.80	36.90	55.33	50.33	36.27	48.15	37.30	37.25	
Read		M	27.38	44.32	35.24	33.10	39.06	37.38	29.62	37.43	36.33	32.67	
	RA	W	31.77	55.22	45.11	37.77	61.49	48.37	36.27	48.65	38.65	35.64	
		M	29.83	41.75	38.32	34.13	44.31	39.59	32.54	40.30	36.44	32.79	
	S	W		4.64	6.96	5.36	6.42	11.20	5.53	9.32	5.01	4.92	
		M		5.04	7.37	6.87	6.42	8.01	6.72	9.23	7.60	6.60	
BSF	R	W		5.06	7.30	5.58	6.23	10.26	5.67	8.96	6.07	5.14	
Read		M		6.10	8.48	6.76	6.56	8.81	6.64	9.91	7.39	6.17	
	RA	W		4.93	6.95	5.73	6.38	8.93	5.79	8.41	6.33	4.95	
		M		5.58	8.76	7.04	6.19	10.22	6.66	9.30	6.73	6.67	
	S	W	48.19	44.44	47.06	39.73	53.67	58.93	36.32	54.95	38.25	39.82	
		M	50.44	38.34	37.18	37.10	39.57	36.86	30.72	41.12	32.20	38.64	
SAT	R	W	45.55	40.39	42.98	37.97	43.20	50.50	35.84	44.07	42.83	36.58	
Math		M	50.10	38.11	37.47	36.74	32.65	36.38	35.36	39.96	36.92	38.92	
	RA	W	44.37	41.83	41.82	39.36	48.72	43.12	35.20	50.08	38.54	39.19	
		M	45.68	38.54	39.93	35.08	38.04	40.54	36.08	43.38	36.23	34.12	
	S	W		5.18	6.60	7.54	8.90	7.07	7.66	9.75	7.46	6.93	
		M		6.12	6.76	9.75	8.64	5.90	9.77	8.94	10.66	9.42	
BSF	R	W		5.52	6.90	8.07	8.51	10.37	7.87	7.27	9.85	6.80	
Math		M		7.63	7.60	9.46	9.85	7.50	9.58	9.87	10.20	8.58	
	RA	W		5.27	6.40	8.18	7.33	11.16	7.89	8.05	9.22	7.04	
		M		6.86	8.34	10.05	8.45	11.32	9.68	8.89	10.13	9.65	

APPENDIX C: Tables of Additional Analyses of STAR Data Pertinent to Small Class as a "Gap Reduction" Strategy for Achievement Differences

Minority Stanford Achievement Test (SAT) Mean Scaled Scores by Grade by Length of Time in Small Class-Size Obtained from "Unfiltered" Data Set

Achievement Measures		Length of Time in Small Class-size			
	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	433.79	_____	_____	_____
	1	498.24	513.74	_____	_____
	2	565.83	565.66	575.28	_____
	3	602.79	606.19	594.55	612.78
SAT Math	K	479.56	_____	_____	_____
	1	511.88	524.42	_____	_____
	2	563.28	562.58	572.39	_____
	3	610.89	604.39	594.90	613.63

White Stanford Achievement Test (SAT) Mean Scaled Scores by Grade by Length of Time in Small Class-Size Obtained from "Unfiltered" Data Set

Achievement Measures		Length of Time in Small Class-size			
	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	443.69	_____	_____	_____
	1	523.60	542.73	_____	_____
	2	587.66	584.93	604.67	_____
	3	611.12	619.10	619.07	633.43
SAT Math	K	496.20	_____	_____	_____
	1	538.08	548.42	_____	_____
	2	582.68	582.42	599.83	_____
	3	615.19	621.32	620.35	635.12

The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT)

Mean Scaled Scores by Grade by Length of Time in Small Class-Size

Obtained from "Unfiltered" Data Set

Achievement Measures	Grade	Length of Time in Small Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	9.90	_____	_____	_____
	1	25.36	28.99	_____	_____
	2	21.83	19.27	29.39	_____
	3	8.33	12.91	24.52	20.65
SAT Math	K	16.64	_____	_____	_____
	1	26.20	24.00	_____	_____
	2	19.40	19.84	27.44	_____
	3	4.30	16.93	25.45	21.49

Minority Small-Class and White Regular-Class Subsample Differences in Stanford Achievement Test (SAT) Mean Scaled Scores by Grade by Length of Time in Class Obtained from "Unfiltered" Data Set

Achievement Measures	Grade	Length of Time in Class 1 year	2 years	3 years	4 years
SAT Reading	K	5.28	_____	_____	_____
	1	18.36	18.96	_____	_____
	2	21.17	14.04	42.22	_____
	3	16.50	12.79	16.06	14.73
SAT Math	K	8.78	_____	_____	_____
	1	16.92	15.58	_____	_____
	2	25.25	16.66	20.39	_____
	3	7.28	19.34	20.83	16.80